Title: The Geodatabase-Implementation of a Utility GIS through the Intranet
Authors: Lance M. Babbitt and Peter Godfrey Jr.

Abstract

In 1999, the Lehigh County Authority, water and wastewater utility located in Allentown, PA, initiated a study to determine the benefits of a geographic information system. After the assembly of a Needs Assessment and Implementation Plan, a conversion pilot and application prototype exercise, the full conversion of all assets into a geodatabase, this past year LCA implemented its enterprise-wide GIS through a web-browser. The configuration is defined by the geodatabase stored in Microsoft SQL Server (LCA DBMS standard), and the established GIS standard platform components (ESRI ArcGIS standard), ArcSDE, and ArcIMS or ArcView/ArcInfo on the desktop. This paper will discuss the GIS environment specifications and configuration, management of both geodatabase and shapefile formatted data, and lessons learned during the implementation. Of particular importance will be a focus on critical aspects of the implementation involving data back up, modifications of map documents, imports of shapefiles into the geodatabase format, and user documentation for management of all aspects of the applications.

Paper Body

Lehigh County Authority (LCA) is a regional provider of water and wastewater services in Lehigh County, Pennsylvania with offices in Allentown. LCA currently supplies 5.2 million gallons per day of water to 14,000 residential, commercial and industrial water customers. We also transport 5.9 million gallons per day of wastewater, primarily through a regional interceptor system. Our service area covers approximately 50 square miles. LCA's 32 full-time employees work in four departments: Administrative, Capital Works, Customer Service and Operations. For several years, LCA has been in a service expansion mode due to high population growth in Lehigh County and an ongoing acquisition program.

Beginning in 1999, LCA undertook efforts to determine if geographic information system technology (GIS) could be integrated into its work processes to increase efficiency. A thorough Needs Analysis and Implementation Plan was completed and identified five (5) priority applications. This effort was followed by a pilot effort to determine and verify the geodatabase design and as well as priority application prototypes were developed for use and comment throughout the organization. Once these efforts were completed, tested and modifications were made to the conversion methods, full system conversion was initiated, which resulted in the delivery of LCA water and wastewater assets in the MS SQL Server geodatabase format.

The next step, and the one being presented, involved identifying the proper components and configuring the hardware and software to enable the data to be

presented on the desktop to both browser based users as well as thick clients who would be doing more intensive analysis with the data. All LCA staff was identified as primary, secondary, or tertiary users (ArcInfo, ArcView, ArcIMS respectively). Under this effort, ArcIMS applications were programmed for discrete user classes in order to allow all LCA employees to access and manipulate data in a display that contained data relevant to their business domain.

The user classes are explained below:

The primary users will require direct access to the GIS data residing on a data server. They require access to a fully functional GIS application that can perform all forms of analysis and data maintenance procedures. These staff will be in charge of maintaining and modifying the GIS data set and must have the tools and access to the data to perform these functions. These staff will require the following applications in order to support their work:

- ESRI ArcGIS – ArcInfo level to perform primary data analysis and all data maintenance activities;
- ESRI ArcPress software for use in efficiently plotting large sized plot files from the GIS;
- ESRI ArcPAD Developer
- SQL server running on the data server to store and maintain the Geodatabase files and access to the SQL application to maintain and modify the database structure when necessary;
- ArcSDE software running on a data server computer to translate the GIS database back and forth between SQL server 2000 on the data server and the ArcGIS software and administrative access to the ArcSDE software in order to maintain the system and data set when necessary.

In order to effectively access and use the GIS data and applications the Primary users will require the following computer hardware components:

- High end computer workstation with a minimum of 700 mhz processor and 1 GB of RAM and Windows 2000 or newer operating system;
- Access to a large format (36” width) color plotter on the computer network with a minimum of 128 MB of RAM to print GIS maps.

Secondary Users

The secondary users will require direct access to the GIS data residing on a data server. They require access to many of the viewing, query and mapping functions within their GIS application and can and will perform most of the GIS functions “from scratch” such as making customized queries and maps. These staff will require the following applications in order to support their work:

- ESRI ArcGIS – ArcView level to perform customized data queries and make customized maps;
- ESRI ArcPress software for use in efficiently plotting large sized plot files from the GIS;
- SQL server running on the data server to store and maintain the Geodatabase files (transparent to the user – no administrative or direct access);
- ArcSDE software running on a data server computer to translate the GIS database back and forth between SQL server 2000 on the data server and the ArcGIS software (transparent to the user – no administrative or direct access).

The SQL and ArcSDE software are required to be running in the background and the ArcView application(s) will access the Geodatabase directly to provide the required data to the Secondary users without the user's knowledge of the SQL or ArcSDE background functions.

In order to effectively access and use the GIS data and applications the Secondary users will require the following computer hardware components:

- High end computer with a minimum of 700 mhz processor and 1 GB of RAM and Windows 2000 or newer operating system;
- Access to a large format (36" width) color plotter on the computer network with a minimum of 128 MB of RAM to print GIS maps.

Tertiary Users

The tertiary users will require indirect access to the GIS data residing on a data server. Access for the tertiary users will occur through a web based GIS interface that streamlines and simplifies the GIS functions and provides users easy, efficient access to specific data about the system. Tertiary users will not perform detailed data analysis nor will they produce highly customized map products. Tertiary users will rely on the Primary and in some cases Secondary GIS users within LCA to perform detailed system analysis or map making. Therefore the Tertiary users will require the following applications in order to support their work:

- Access to the GIS data through a customized web based interface using ESRI ArcIMS software;
- SQL server running on the data server to store and maintain the Geodatabase files (transparent to the user – no administrative or direct access);
- ArcSDE software running on a data server computer to translate the GIS database back and forth between SQL server 2000 on the data server and the ArcGIS software (transparent to the user – no administrative or direct access).

The SQL and ArcSDE software are required to be running in the background and the ArcIMS application(s) will access the Geodatabase directly to serve the data to the Tertiary users without the user's knowledge of the SQL or ArcSDE background functions.

After an analysis was done for the tertiary users, it was determined that their PC configuration would not need to be upgraded but, In order to effectively access and use the GIS data and applications the Tertiary users will require the following computer hardware components:

- A computer with Microsoft Internet Explorer version 5.0 or above;
- Access to a dedicated 11" x 17" color ink jet printer on the network in order to print color output from the customized ArcIMS application.

Once the user requirements were defined, the hardware and software requirements were defined as follows:

Hardware

- Dedicated Application Server - One (1) new server dedicated to serving the GIS data to users via the web based ArcIMS applications (Tertiary users). The application server will initially host the ArcIMS application and its specifications will support additional growth in the use of the ArcIMS applications as well as other future GIS-based applications (work order management, asset management, etc);
- Dedicated GIS Data Server – One (1) new GIS server, identical in specifications to the application server, will be purchased to serve as a dedicated data server. The GIS data will be migrated from the application server to the data server to reduce the load on the existing server. Primary and Secondary users will access the data on the server through direct connections while the ArcIMS application on the Application Server will access the data to serve through ArcIMS to the Tertiary users. This server will store the Geodatabase data in MS SQL Server database files and will serve the data to all users through ArcSDE software.
- DLT Tape Backup Unit – One (1) new 40 GB uncompressed/80 GB compressed DLT tape backup unit to backup all data files on the new (and potentially existing) servers;
- High End PCs – Three (3) new computers to be used by the Secondary users in order to enable them to use ArcGIS/ArcView;
- Workstation Computer - One (1) new high end GIS workstation computer for the Primary user in order to better utilize the GIS software and data;
- Large Format Color Plotter – One (1) new, 36" wide color plotter capable of printing large GIS maps from anyone on the network;
- HP iPAQs – Two (2) new PDAs for the piloting and integration of data collection by field staff.

Software

- ESRI ArcIMS Web Server Software – a copy of ESRI's web serving software, ArcIMS, in order to serve the customized web based applications to the Tertiary users;
- ESRI ArcSDE Software – a copy of ESRI's Geodatabase management and translation software, ArcSDE, to manage the Geodatabase in the SQL environment;
- Microsoft SQL Server Standard Edition software – database serving software to store the Geodatabase to all users and the web server software;
- ESRI ArcView Licenses – Four (4) floating licenses of ArcView 8.3 to be used by all secondary users
- ESRI ArcInfo License – one (1) license of ArcInfo 8.3 to be used by the primary user
- ESRI ArcPress for ArcGIS software – a copy of ArcPress for ArcGIS which is used to assist in the plotting of large maps from the GIS and speed the plotting time;
- ESRI ArcPAD and ArcPAD Application Builder – two (2) copies of ArcPAD (LCA already has one copy) and one (1) copy of ArcPAD application builder for field data collection interface development.

The data requirements included both internal LCA asset information as well as data that was maintained and stored in Lehigh County on their mainframe. This aspect required procedures to be piloted in order for LCA to efficiently obtain data that included tax parcels, street centerlines, municipal and county boundaries, orthophotography, stored in shapefiles, and customer account data stored in ADMINS format.

LCA's data was stored in a geodatabase and accessed through SDE. Data on the County mainframe was stored in both shapefiles and the previously mentioned ADMINS format. Procedures were put into place to transfer county shapefiles on a cyclical timeframe to LCA and it would then be imported into a geodatabase format and the SDE connections and ArcIMS services would be updated. Initially this transfer was by CD. Currently, it is accessed through an FTP site with the County.

The ADMINS data needed to be transferred in a semi-colon delimited format and then imported into a SQL server table. Geocoding this data to existing parcel centroids created in the County shapefiles provided functionality to link the created meter centroids to an associated water main in order to provide a quick analysis of customers who may be impacted if there was a water main break (one of the priority applications) as well as a general query on customer account information. All the data needed to be indexed and configured in a revised ArcView .mxd document for service over the browser. Each time new data is acquired, the service needs to be updated and minor maintenance procedures

need to be run to integrate any new meters as well as modify existing ones that have been changed.

The LCA implemented environment consisted of two DELL servers GISDATA (holding MS SQL Server and ArcSDE) and GISAPP (running the ArcIMS service). When the configuration was complete, the browser was tested across the enterprise for completeness and functionality. The browser had been configured with four (4) tabs representing the intended client base (capital planning, operations, customer service, and administration) and was launched in September 2003. Each tab is linked to its own mxd file. After being trained, all users were given a period of time to interact with the site and request modifications. Some of the modifications could be made in the ArcIMS service, while others required an assignment of a thick client to the user based upon their requirements. As this paper is being written, LCA is still in the process of determining true user base needs and modifying the service to meet these needs.

Over the next year, modifications and enhancements are planned for the LCA GIS. These include finalizing procedures, piloting, and automating external data imports, conversion, and integration with the dataset; modifications to the web service based upon user input and extension of the user base to include some internal sectors that have not fully taken advantage of the new tool.

The implementation of GIS across an enterprise offers many chances to increase efficiency, but at the same time, overlook some aspects of implementation that need to be considered. Lessons learned would be a logical name. In this effort to date, the following list contains lessons learned:

- Never overestimate the level of technical understanding of the intended audience. The integration of a new tool into a legacy process creates confusion regardless of the training level. The user base is the measurement of success of any technical implementation and their level of understanding needs to be considered and user interface options require full understanding

- The investment in data requires that adequate measures be made to “back-up” all data on a prescribed timeframe. The back-up procedures, while sounding trivial, actually required a large level of effort to work out the “kinks” and get the procedure streamlined. For instance, to obtain a full system backup, done weekly, the entire ArcIMS web service needed to be shut down before back up and restarted after.

- As is typical with large implementations, there is a requirement to not only manage and update internal data, but also integrate with external data that is maintained and updated by others. This data can also be maintained in a different format and on a different schedule, which requires additional effort to integrate with an application dataset. In the case of LCA, there are numerous

required “base data layers” that are needed and they are maintained by another entity in a different format. The requirements to update the application data often needs to have discrete procedures to minimize error as well as facilitate currency

- The importance of full documentation cannot be over-estimated. Every aspect of the process from the specifications of the users through the data transfer methodologies has been documented. By doing this an accountability thread has been created as well as the potential for LCA to continue its discovery and growth of GIS internally with minimal external assistance.

As LCA’s GIS moves into future years, the emphasis will be on keeping current with any software upgrades and integration requirements, continue to increase the user base through additional application development, and the development of an external customer website.

Author(s) Information

Primary Author: Lance M. Babbitt
GIS Project Manager
Lehigh County Authority
P.O. Box 3348
1053 Spruce Street
Allentown, PA 18106-0348
610.398.2503
lancebabbitt@lehighcountyauthority.org

Secondary Author: Peter Godfrey Jr.
CDM
1500 JFK Boulevard, Suite 624
Philadelphia, PA 19102
215.636.0600 EXT. 254
godfreyp@cdm.com